

Dr. Torsten Beyer

Nachhaltige Websites: Datenvolumen und CO₂-Emissionen ermitteln und überwachen

Jede Website trägt durch ihr stetig steigendes Datenvolumen dazu bei, dass der Stromverbrauch des Internets und damit die CO₂-Emissionen zunehmen. Nach seiner grundlegenden Einführung in das Thema in der WebsiteBoosting #75 zeigt der Experte für Web-Nachhaltigkeit Dr. Torsten Beyer, wie der Aufbau eines Monitorings für die Nachhaltigkeit von Websites gelingt und mit welchen Metriken sie beschrieben und überwacht werden kann. Durch die Analyse einzelner Unterseiten mit CO₂-Kalkulatoren oder der Browser-Konsole können keine belastbaren Zahlen gewonnen werden, das gelingt nur durch eine Analyse der Server-Logfiles. Auch mit Blick auf die Bilanzierung von CO₂-Emissionen lohnt sich eine Beschäftigung mit diesem aktuellen Thema.

Entwicklung des Datenvolumens

Das durchschnittliche Datenvolumen einer Website betrug 2022 laut Web Almanac 2,3 Megabyte für die Desktop-Version und 2,0 Megabyte für die mobile Version. Vor zehn Jahren lagen die Werte noch bei 750 (Desktop) bzw. 300 (mobil) Kilobyte. Dieser enorme Zuwachs ist der Tatsache geschuldet, dass Internetzugänge in dieser Zeit immer schneller wurden und man sich kaum noch Gedanken um das übertragene Datenvolumen und langsame Ladezeiten machen musste. Ein typischer Rebound-Effekt also, der allerdings auf Kosten der Natur geht. Denn mehr Datenübertragungen bedeuten einen höheren Stromverbrauch und in letzter Konsequenz mehr CO₂-Emissionen. Diese Erkenntnis ist bei vielen noch nicht angekommen und eine Trendwende ist nicht absehbar.

Die in den vergangenen Jahren sukzessive eingeführten Core Web Vitals und das „Mobile First“-Update waren ein eher halbherziger Schritt von Google, um einen schnelleren, stabilen Seitenaufbau mit möglichst unmittelbarer Interaktionsmöglichkeit zu belohnen – und damit indirekt auch ein geringeres Datenvolumen. Da sie allerdings bislang nur ein sehr

schwaches Ranking-Kriterium sind, verwundert es nicht, dass laut Web Almanac bislang nur knapp 40 % aller Websites die Vorgaben erfüllen. Daher hat dieses Thema bei der Suchmaschinenoptimierung nur eine niedrige Priorität.

Nachhaltigkeit ist (noch) kein Ranking-Faktor

Wer sich um eine nachhaltigere Website bemüht, muss das primär aus einer intrinsischen Motivation heraus tun, etwas gegen den Klimawandel unternehmen zu wollen. Oder sieht den sparsamen Umgang mit Ressourcen als Teil seiner unternehmerischen Verantwortung an. Das kann heute in der Kundenkommunikation vorteilhaft sein, denn Seitenbesucher haben kaum Einfluss auf das generierte Datenvolumen, sie können höchstens durch ihr Nutzungsverhalten nachhaltigere Seiten bevorzugen. Einige Betreiber wollen ihre Website vielleicht auch nur mit einem Nachhaltigkeitssiegel aufwerten. Aber Vorsicht: Es gibt eine Vielzahl von Anbietern und bisher keine echte Zertifizierung, daher ist die Gefahr von Greenwashing hoch – insbesondere wenn keinerlei Anpassung an den Seiten erforderlich ist und man lediglich ein paar

DER AUTOR



Dr. Torsten Beyer ist Chemiker und Gründer der Dr. Beyer-Internet-Beratung. Seit über 25 Jahren optimiert er Websites für schnelle Ladezeiten. Und seit Kurzem auch für einen minimalen digitalen Fußabdruck.

Euro für oft zweifelhafte CO₂-Zertifikate investieren muss.

Alle, die eine ernsthafte Absicht zum Wohl der Natur verfolgen, verdienen Respekt. Das Sustainable Web Manifesto aus dem Jahr 2019 liefert fundierte Handlungsempfehlungen für die Gestaltung nachhaltiger Websites (<https://www.sustainablewebmanifesto.com/>). Es wurde bisher von knapp 2.800 Menschen weltweit unterzeichnet.

Ein Megabyte pro Seite geht!

Es ist übrigens durchaus möglich, mit einem durchschnittlichen Datenvolumen von einem Megabyte pro Seite optisch ansprechende Websites zu erstellen. Daher empfiehlt es sich, einen Schwellenwert für die eigene Website zu definieren und das Design und die technische Optimierung danach auszurichten. Dazu genügt in der Regel die Nutzung moderner Bild- und Schriftformate, der Einsatz von „lazy loading“, eine vernünftige responsive Version für mobile Endgeräte und das Abrufen überflüssiger JavaScript-Bibliotheken, um die wichtigsten Schrauben zu nennen.

Probleme bei der Ermittlung des Datenvolumens

Die in der WebsiteBoosting #75 vorgestellten Tools zur Abschätzung der CO₂-Emissionen und des Datenvolumens von Websites (Website Carbon Calculator, Digital Beacon und Ecograder) haben mehrere Nachteile: Die Hochrechnung auf die gesamte Website beruht immer nur auf der Analyse einer einzelnen Unterseite, in der Regel der Startseite. Alle Prognosen aufgrund der Besucherzahl und der Seitenaufrufe pro Besucher für einen größeren Zeitraum, beispielsweise ein Jahr, können daher sehr fehlerhaft sein und sind keine Basis für den Aufbau eines Monitorings für den Erfolg einer nachhaltigen Opti-

mierung einer kompletten Domain.

Natürlich sind die Tools sehr hilfreich, um grundlegende templatebasierte Probleme mit Schriften, JavaScript, CSS, Header- und Footer-Bildern oder im HTML-Code zu finden. Individuelle Probleme auf einzelnen Unterseiten mit sehr vielen oder nicht optimierten Bildern können damit aber ebenso wenig gefunden werden wie große PDF-Dateien oder lokal gehostete Videos. Und der Effekt jedes einzelnen Problems ist natürlich abhängig von den individuellen Zugriffszahlen der betroffenen Seiten.

Zusammenfassend kann man sagen, dass die durch die Tools gewonnenen Daten hilfreich für die Optimierung von Seitentemplates sind, allerdings taugen sie wenig für eine Gesamtbilanzierung. Google Lighthouse (Labordaten) und die Google Search Console (Felddaten) mit ihrer unterschiedlichen Datenbasis lassen grüßen.

Labordaten sind keine Felddaten!

Um echte Felddaten zu erhalten, müssen das verwendete Endgerät, der genutzte Browser und die Bildschirmauflösung des Betrachters ebenso einbezogen werden wie die Verweildauer auf einer Seite, die Zahl der Seitenaufrufe pro Besuch und wie weit jeweils nach unten gescrollt wird. Denn all diese Faktoren beeinflussen das übertragene Datenvolumen unter Umständen massiv. Auch wer schon vor dem vollständigen Laden der Seite weiterravigiert, generiert ein ganz anderes Datenvolumen als jemand, der länger auf einer Seite verweilt und vielleicht ganz nach unten scrollt. Denn ist das „lazy loading“ richtig konfiguriert, dann werden audio-visuellen Elemente erst geladen, wenn sie in den sichtbaren Bereich („Viewport“) des Betrachters kommen. Ebenso kann das individuell konfigurierte Caching von Inhalten einen großen Einfluss für wiederkeh-

INFO

Eigentlich müsste man wissenschaftlich korrekt immer von CO₂-Äquivalenten (kurz: CO₂eq) sprechen, da Kohlenstoffdioxid nur eines von vielen Treibhausgasen ist, die in einer Kennzahl anteilig zusammengefasst werden. Da im allgemeinen Sprachgebrauch aber überwiegend von Emissionen gesprochen wird, wird hier ausschließlich diese Formulierung verwendet. Gemeint sind also immer die Äquivalente.

rende Besucher haben, die einzelne Elemente noch im Browser-Cache haben. Für die Ladezeiten und für die Core Web Vitals sind solche Optimierungen im Übrigen immer sinnvoll. Ein nachhaltige Website berücksichtigt all diese Faktoren und der Server liefert immer die datensparsamste Variante in Abhängigkeit vom Endgerät aus. Das geht über eine reine Optimierung der Ladezeiten hinaus, denn die lässt sich auch über leistungsfähigere Webserver, die Auslagerung von Daten in Content-Delivery-Netzwerke oder JavaScript-Tricksereien erreichen.

Darüber hinaus spielt auch die serverseitige Komprimierung von Textdateien (HTML, XML, CSS, JS, SVG) eine Rolle. Die meisten modernen Browser (aktuell mehr als 96 %) verstehen inzwischen die effizientere BROTLI-Komprimierung. Dadurch werden weniger Daten übertragen als bei älteren, die nur die DEFLATE-Komprimierung verstehen. Wir reden hier in der Regel von einem Einsparpotenzial im niedrigen zweistelligen Prozentbereich. Es gibt auch immer noch Server, auf denen gar keine Komprimierung aktiviert wurde, was das übertragene Datenvolumen massiv erhöht. Auch viele Crawler und Bots nutzen übrigens nur die DEFLATE-Komprimierung.

Auf Mobilgeräten werden oft andere Bildgrößen geladen als auf großen Desktopmonitoren. Und wer moderne

```

79.233.164.80 - - [24/Apr/2023:16:41:47 +0200] "GET / HTTP/2.0" 200 8009 "-" "Mozilla/5.0 ..."

Weitere Einträge ohne IP-Adresse, Zeitstempel, Referrer und User-Agent:

"GET /grafik/header.webp HTTP/2.0" 200 12767
"GET /analytik.css HTTP/2.0" 200 5436
"GET /grafik/analytiknews.svg HTTP/2.0" 200 2355
"GET /analytik.js HTTP/2.0" 200 1173
"GET /grafik/labor.svg HTTP/2.0" 200 6056
"GET /m/matomo.js HTTP/2.0" 200 21630
"GET /ad/www/delivery/asyncjs.php HTTP/2.0" 200 2699
"GET /grafik/sl1300401.webp HTTP/2.0" 200 12588
"GET /grafik/envelope.webp HTTP/2.0" 200 3330
"GET /favicon.png HTTP/2.0" 200 1604

```

Abb. 1: Insgesamt elf Hits beim Aufruf der Startseite einer Domain mit einem Datenvolumen von insgesamt 77.647 Kilobyte (Formate HTML, PHP, CSS, JS, SVG, WEBP und PNG). Hat ein Nutzer Inhalte im Cache, werden nicht alle Dateien geladen. Bei einem älteren Browser mit DEFLATE-Komprimierung wäre das Datenvolumen bei Textdateien deutlich höher. Bei mobilen Zugriffen und responsiver Darstellung würden unter Umständen andere Bilddateien geladen.

Bildformate wie WEBP oder AVIF einsetzt, wird in der Regel noch eine Fallback-Lösung auf JPEG oder PNG-Varianten vorhalten, deren Dateigröße deutlich höher sein wird. Das Gleiche gilt für Fallback-Lösungen für das Schriftformat WOFF2, beispielsweise auf WOFF, TTF oder OTF. Es gibt noch mehr Szenarien (Browserweichen, Cookies, Skripte), die das Datenvolumen individuell beeinflussen.

Und als ob das noch nicht genug Probleme wären, haben aktuell alle Testtools ein Problem mit DSGVO-konformen Cookie-Consent-Bannern, weil sie die nachgeladenen Daten bei einer Zustimmung nicht erfassen und damit deutlich zu niedrige Datenvolumina suggerieren, die nicht der Realität entsprechen. Auch das „lazy loading“ können sie nicht berücksichtigen, weil sie nicht scrollen können.

Helfen Analytics-Tools bei der Ermittlung des „echten“ Datenvolumens?

Leider nicht, denn damit werden ja nur Pageimpressions und Besucherzahlen erfasst und nicht alle anderen Seitenelemente. Und insbesondere PDF-Dateien und lokal gehostete Video- und Audiodateien können richtige Datenmonster sein, die bei einer solchen Betrachtung komplett durchs Raster fallen. Und nur wenn mit der verwendeten Cookie-Consent-Lösung die Ableh-

nungsquote ermittelt werden kann, hat man die echten Besucherzahlen und die Seitenzahl pro Besuch, die für die spätere Auswertung wichtig sind. Ansonsten kann man die Ablehnungsquote aus dem Schwund der Zugriffe in Google Analytics gegenüber den Zahlen aus der Google Search Console grob abschätzen. Wer auch diese Daten nicht hat, kann mit einer Zustimmungsrate von durchschnittlich 40 % rechnen (*ein-fach.st/cookief7*). Allerdings sind diese Zahlen mit Vorsicht zu genießen, da sie je nach der Gestaltung des Cookie-Banners stark variieren können.

Was ist mit Crawlern wie dem Screaming Frog?

Solche Tools sind sehr nützlich, um große Dateien (z. B. PDF-Dokumente, Bilder, Schriften, Videos) oder veraltete Dateiformate auf einer Website zu finden und dann entsprechend zu optimieren. Verknüpft man den Screaming Frog mit der PageSpeed API von Google, dann kann sogar ein theoretisches Datenvolumen für jede Datei ermittelt werden. Allerdings simuliert der Frog standardmäßig die DEFLATE-Komprimierung, die bei realen Zugriffen und von Bots aber nur teilweise verwendet wird. Man kann das zwar standardmäßig auf BROTLI-Komprimierung ändern, allerdings lässt sich auch damit das reale Geschehen nicht abbilden. Genauso wenig wie „lazy loading“,

Caching, Fallback-Lösungen (Bilder, Schriften, Videos) für ältere Datenformate oder das Laden anderer Bildgrößen auf kleinen Bildschirmen. Kurzum: Crawler helfen bei der Diagnose und der Ermittlung von Verbesserungsmöglichkeiten. Felddaten für das reale übertragene Datenvolumen können damit aber ebenso wenig ermittelt werden wie mit Analytics-Tools oder den Tools zur Analyse einzelner Seiten.

Die Lösung: Server-Logfiles!

Der heilige Gral und damit die Lösung all dieser Probleme sind die Access-Logfiles des eigenen Webserver bzw. Webspace beim Provider. In den 1990er-Jahren stellten sie noch die einzige Möglichkeit zur Auswertung der Zugriffe auf eine Website dar, sind aber durch die Verbreitung von Analytics-Tools wie Urchin seit 1998, das später von Google übernommen und in Google Analytics umbenannt wurde, heute weitgehend in Vergessenheit geraten. Das liegt auch daran, dass Logfiles relativ unhandlich und groß werden können, weil für jeden Dateizugriff („Hit“) eine eigene Protokollzeile geschrieben wird. Nicht selten können so bei einer Pageimpression Dutzende Hits entstehen, wenn Bilder, Schriften, JavaScript- oder CSS-Dateien zur Darstellung geladen werden müssen.

Die Einträge in den Server-Logfiles wirken auf den ersten Blick etwas

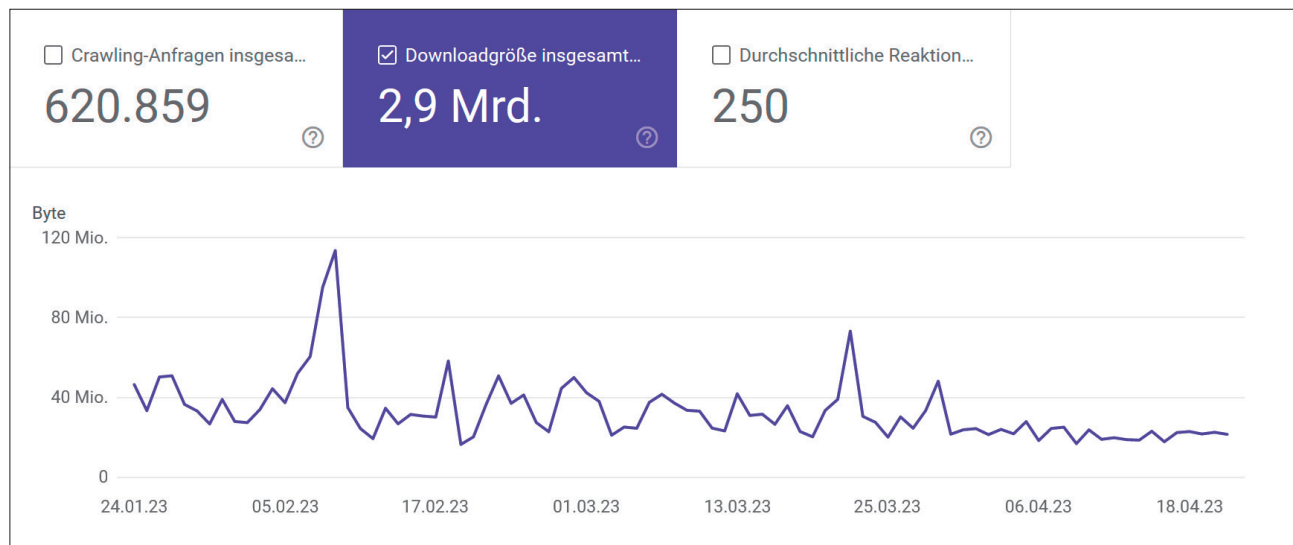


Abb. 2: In der Google Search Console können für jede Domain die Zahl der Anfragen und das generierte Datenvolumen beim Crawl in den letzten 90 Tagen eingesehen werden unter „Einstellungen“ -> „Crawling-Statistiken“. Im Beispiel generierten knapp 621.000 Anfragen ein Datenvolumen von 2,9 Gigabyte (= 2,9 Milliarden Byte).

kryptisch. Die entscheidende Zahl, das übertragene Datenvolumen in Byte, ist daher in Abbildung 1 fett gedruckt. Der große Vorteil: Das ist der reale Wert für die angeforderte Datei unter Berücksichtigung der serverseitigen Komprimierung und der individuellen Einstellungen für den Browser des Endgeräts. Werden Dateien aus dem Cache oder bei korrekt konfiguriertem „lazy loading“ gar nicht geladen, dann tauchen sie im Logfile nicht auf. Und bei allen Fallback-Lösungen in Abhängigkeit vom Browser des Endgeräts und der Bildschirmgröße steht im Logfile die URL der wirklich geladenen Datei.

Bot-Traffic kann überwiegen

Server-Logfiles haben noch einen weiteren entscheidenden Vorteil. Nur hier sieht man alle automatisierten Zugriffe von Crawlern wie dem Google-Bot oder neuerdings ChatGPT, ebenso

wie alle Zugriffe von SEO-Tools und allen zwielichtigen Bots, die Daten abgreifen, Sicherheitslücken suchen oder schlimmstenfalls eine DDoS-Attacke fahren wollen. Es gibt Schätzungen, dass der gesamte Bot-Traffic über 42 % des gesamten generierten Datenvolumens einer Domain ausmacht. Berücksichtigt man die durchschnittliche Zustimmungsquote von Cookie-Consent-Bannern (40 %), dann müssen alle mit Analytics-Tools ermittelten Werte für das Datenvolumen mit sechs multipliziert werden, in Extremfällen könnte es sogar ein Faktor von 100 sein, wenn eine Domain 90 % Bot-Traffic hat und 90 % dem Cookie-Consent widersprechen würde. Für eine ehrliche Bilanz des gesamten Datenvolumens einer Domain gehören beide Faktoren dazu, sonst kann der Rechenfehler kolossal sein! Den Bot-Traffic zu optimieren, ist eine Wissenschaft für sich, zählt aber aus den genannten Gründen genauso

zu einer nachhaltigen Optimierung dazu.

Wenn wir das Bild eines Eisbergs bemühen, dann sieht ein Analytics-Tool nur einen Teil der Oberfläche und das Logfile alles auch unter der Oberfläche. Wer sehen will, wie viel Datenvolumen beispielsweise der Google-Bot generiert, der findet dazu eine Auswertung der letzten 90 Tage in der Search Console unter „Einstellungen“ -> „Crawling-Statistiken“ (Abbildung 2) und wird vielleicht überrascht sein. Mit effizienterem Crawling könnte Google übrigens auch viel für die Umwelt tun und nebenbei Geld sparen.

Auswertungsmöglichkeiten für Logfiles

Am einfachsten nutzt man zur Auswertung die vom Provider angebotene Software im Kunden-Backend, die im Idealfall mehrere Jahre, manchmal aber auch nur einen Monat zurückreicht.

Zusammenfassung					
Zeitraum	Monat März 2023				
Erster Zugriff	01.03.2023 - 00:00				
Letzter Zugriff	31.03.2023 - 23:59				
	Unterschiedliche Besucher	Anzahl der Besuche	Seiten	Zugriffe	Bytes
gesehener Traffic *	46.635	60.431 (1.29 Besuche/Besucher)	433.359 (7.17 Seiten/Besuch)	956.232 (15.82 Zugriffe/Besuch)	8.09 GB (140.3 KB/Besuch)
nicht gesehener Traffic *			1.188.342	1.552.178	12.40 GB

Abb. 3: Eine beispielhafte Auswertung für einen Monat mit dem Tool AWStats im Kunden-Backend des Providers Hetzner. Die Domain hat über 60 % Bot-Traffic („nicht gesehener Traffic“). Die weiteren Metriken können zum Aufbau eines monatlichen Monitorings verwendet werden.

Zum Aufbau eines Monitorings ist eine monatliche Auswertung durchaus sinnvoll. Später kann man natürlich auch einen wöchentlichen Report und bei einem Relaunch oder größeren Arbeiten an der Seite eine tägliche Auswertung etablieren.

Wer diesen wertvollen Datenpool uneingeschränkt nutzen will, der sollte auf seinem Webspace in einem passwortgeschützten Bereich eine eigene Auswertungssoftware installieren und individuell konfigurieren, wenn ihm die angebotene Auswertung des Providers nicht ausreicht. Alternativ können die Logfiles natürlich auch heruntergeladen und lokal ausgewertet werden. Das generiert bei großen Seiten allerdings viele Gigabyte Datenübertragungen, was dann auch wieder nicht nachhaltig ist. Für eine lokale Auswertung gibt es eine Vielzahl kostenfreier und kostenpflichtiger Tools. Oft steht dort aber das Datenvolumen nicht im Mittelpunkt, sondern das Finden von Fehlern oder die Crawlsteuerung wie beispielsweise beim Screaming Frog Logfile Analyser. Eigentlich reicht ein Tool wie AWStats (awstats.sourceforge.io) für das angestrebte Nachhaltigkeitsmonitoring völlig aus. Natürlich kann man auch mit Excel für kleinere Seiten oder Tools wie KNIME für große Domains eigene Auswertungen aufsetzen.

Es kann leider vorkommen, dass ein direkter Zugriff auf die Logfiles im aktuellen Hosting-Paket nicht möglich ist. Teilweise werden auch die Daten aller Domains, die man bei einem Provider betreibt, in einem File zusammengefasst. Hier lohnt eine Nachfrage, ob ein Upgrade in ein anderes Paket Abhilfe schafft. Oder man definiert diese Anforderung für den nächsten Relaunch, der vielleicht mit einem Host-Wechsel verknüpft werden kann. Denn die Auswertung dieser Daten könnte im Zuge einer CO₂-Bilanzierung irgendwann auch Websites betreffen, und dann ist man vorbereitet.

Datei-Typen					
	Datei-Typen	Zugriffe	Prozent	Bytes	Prozent
	php Dynamic PHP Script file	274.279	29.1 %	1.73 GB	20.4 %
	html HTML or XML static page	157.653	16.7 %	1.75 GB	20.6 %
	xml HTML or XML static page	104.443	11.1 %	517.07 MB	5.9 %
	jpg Image	101.186	10.7 %	1.46 GB	17.2 %
	js JavaScript file	79.745	8.4 %	548.27 MB	6.3 %
	svg Image	64.028	6.8 %	237.49 MB	2.7 %
	png Image	60.731	6.4 %	399.86 MB	4.6 %
	gif Image	40.936	4.3 %	226.90 MB	2.6 %
	css Cascading Style Sheet file	27.320	2.9 %	136.88 MB	1.5 %
	webp Image	25.982	2.7 %	483.05 MB	5.5 %
	pdf Adobe Acrobat file	2.367	0.2 %	1013.67 MB	11.6 %
	gz Archive	1.053	0.1 %	42.60 MB	0.4 %
	txt Text file	129	0 %	664.00 KB	0 %

Abb. 4: Beispielhafte Aufschlüsselung der Zahl der Zugriffe und des generierten Datenvolumens auf die verwendeten Datei-Formate mit AWStats. Die Website nutzt schon moderne Bildformate (WEBP, SVG), hat aber noch viele Bilder in alten Formaten (JPG, PNG, GIF), die umgewandelt werden sollten. PDF-Dateien machen zwar nur 0,2 % der Zugriffe, aber 11,6 % des gesamten Datenvolumens aus.

Sind diese Fragen geklärt, sollte man aus den gewonnenen Daten ein Monitoring aufsetzen und eigene Metriken definieren. Dabei gehen wir im Folgenden davon aus, dass die Daten im Kunden-Backend des Providers dafür ausreichen. Aus einer monatlichen Auswertung lassen sich verschiedene Metriken ableiten:

- » Der Anteil des Bot-Traffics am gesamten Datenvolumen
- » Das generierte Datenvolumen pro Zugriff, Seitenaufruf oder Besucher
- » Die Betrachtung inklusive oder exklusiv des Bot-Traffics

Prinzipiell empfiehlt es sich, den Bot-Traffic gesondert zu betrachten. Denn eine Optimierung erfordert andere Methoden als die Optimierung des Besucher-Traffics, auf den wir uns hier fokussieren. Im Beispiel generiert ein Besucher ein Datenvolumen von durchschnittlich 173 Kilobyte. Das ist ein sehr guter Wert, wenn wir die eingangs genannten 2,3 Megabyte heranziehen. In diesem einen Wert akkumulieren sich alle Optimierungen des Datenvolumens, also eine Art „nachhaltiger PageRank“, der unabhängig von Provider und vom verwendeten Content-Management-System ist und Websites untereinander vergleichbar macht. Um alle saisonalen Effekte zu bereinigen, sollte der Vergleichszeit-

raum auf ein komplettes Jahr ausgeweitet werden. Da der Bot-Traffic andere Ursachen hat, sollte man ihn getrennt betrachten und optimieren.

Auch die Zahl der Zugriffe pro Besuch ist übrigens eine sinnvolle Metrik, um Schwankungen der Besucherzahlen zu berücksichtigen. Denn wenn man beispielsweise die Zahl der JavaScript- und CSS-Dateien reduziert oder Logos in SVG-Sprites zusammenfasst, wird sich das in dieser Zahl manifestieren.

Den maximalen Nutzen aus der Logfile-Auswertung zieht man durch eine Aufschlüsselung des gesamten Datenvolumens auf die verwendeten Datei-Formate (Abbildung 4). Jede Bildoptimierung, jede Änderung der Caching-Regeln, jede neue Fallback-Lösung für ältere Browser, jede responsive Optimierung, eine Umstellung auf die serverseitige Brotli-Komprimierung etc. lässt sich hier ablesen. Wie die eigene Monitoring-Lösung letztlich auch aussieht, die Auswertung der Server-Logfiles liefert dazu umfangreiches Zahlenmaterial.

Externe Inhalte

Wie man mit eingebundenen Dateien von anderen Quellen umgeht, ist noch einmal ein ganz anderes Thema. Dabei reden wir über eingebun-

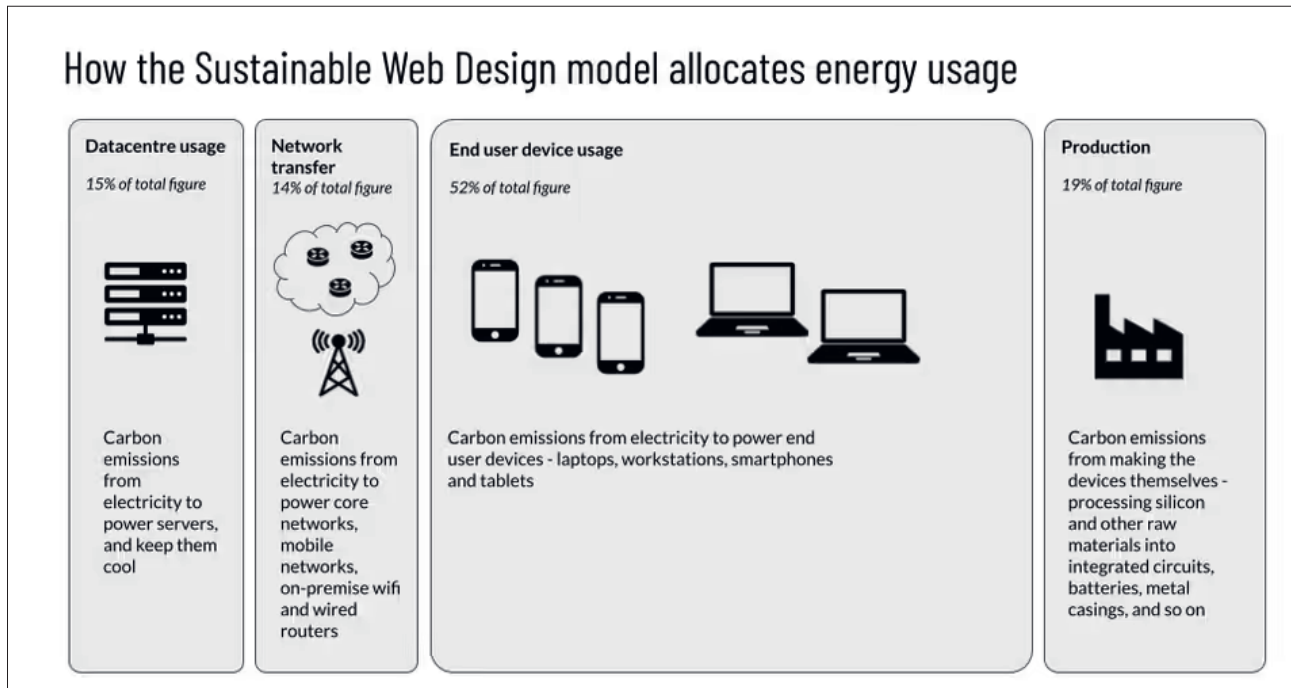


Abb. 5: Das Sustainable Web Design Model (SWD). Danach teilt sich der Stromverbrauch beim Betrachten einer Website auf vier Bereiche auf. Das Rechenzentrum des Providers schlägt dabei mit 15 % zu Buche, der Datentransfer zwischen Endgerät und Rechenzentrum mit 14 %, die Herstellung der Endgeräte mit 19 % und der Stromverbrauch auf den Endgeräten macht mit 52 % den Löwenanteil aus.

dene Video- und Audiodateien, über die Player entsprechender Plattformen oder alles, was an statischen Elementen auf ein Content Delivery Network (CDN) ausgelagert wird. Und natürlich Cookies. Teilweise sind bei den Dienstleistern eigene Auswertungen verfügbar, die dann natürlich ins Monitoring einbezogen werden sollten. Leider erschweren externe Inhalte das Monitoring des Datenvolumens. Es sollte aber klar sein, dass die eigene Bilanz nicht dadurch aufgehübscht werden kann, indem man statische Dateien in ein CDN auslagert, die dann nicht mehr in den eigenen Logfiles auftauchen.

Wie kommt man vom Datenvolumen zu CO₂-Emissionen?

Während das gesamte Datenvolumen einer Domain bis auf die genannten Einschränkungen bei eingebundenen Elementen im Prinzip bis aufs Byte genau über die Logfiles ermittelt werden kann, ist man bei der Abschätzung der resultierenden CO₂-Emissionen auf Schätzungen und vereinfachte Annahmen angewiesen.

Von den meisten Kalkulatoren wird das sogenannte Sustainable Web Design Model (SWD) angewendet (Abbildung 5). Natürlich dürfen die dort genannten Beiträge nicht isoliert betrachtet werden, denn eine schwere Website mit viel Datenvolumen erfordert wahrscheinlich zudem leistungstärkere Hardware und mehr Speicherplatz. Sie verbraucht auch mehr Strom am Endgerät, wenn der Prozessor übermäßig mit Rendering von JavaScript und komplizierten CSS-Verschachtelungen beschäftigt ist, und führt beim Endnutzer wegen der höheren Anforderungen vielleicht dazu, dass er sich genötigt sieht, ein neues Endgerät mit mehr Speicher oder einem schnelleren mobilen Netzzugang anzuschaffen. Das wäre dann wieder der gefürchtete Rebound-Effekt, der schlussendlich zu immer mehr Strom- und Ressourcenverbrauch und in letzter Konsequenz zu mehr CO₂-Emissionen führt. Kritiker fügen dann gerne an, dass die Einzelbeiträge kaum messbar sind, aber aktuell sind über 5 Milliarden Menschen im Internet unterwegs. Viele Schneeflocken sind an sich kein

Problem, aber wenn sich zu viele in die gleiche (falsche) Richtung bewegen, entstehen verheerende Lawinen. Daten sind zwar unsichtbar, aber in letzter Konsequenz passiert das Gleiche.

Abgeleitet von diesem Modell stellt die Green Web Foundation über die JavaScript-Bibliothek CO₂.js umfangreiches Zahlenmaterial zur sogenannten „Grid Intensität“ einzelner Länder, Kontinente oder einen globalen Mittelwert zur Verfügung (developers.thegreenwebfoundation.org). Darunter versteht man die CO₂-Emissionen, die pro erzeugter Kilowattstunde Strom entstehen. Der Wert wird umso kleiner, je weniger fossile Kraftstoffe für die Stromproduktion eingesetzt werden. Und ja, Atomstrom erzeugt auch fast keine CO₂-Emissionen, ist aber wegen der ungelösten Entsorgungsfrage in Deutschland Geschichte. In Statistiken wird sein Anteil nicht explizit genannt, sondern ergibt sich aus der Differenz des Anteils von regenerativem Strom und CO₂-armem Strom (Abbildung 6). Das sollte man bedenken, wenn man vielleicht einen Providerwechsel nach Frankreich oder Norwegen plant, weil

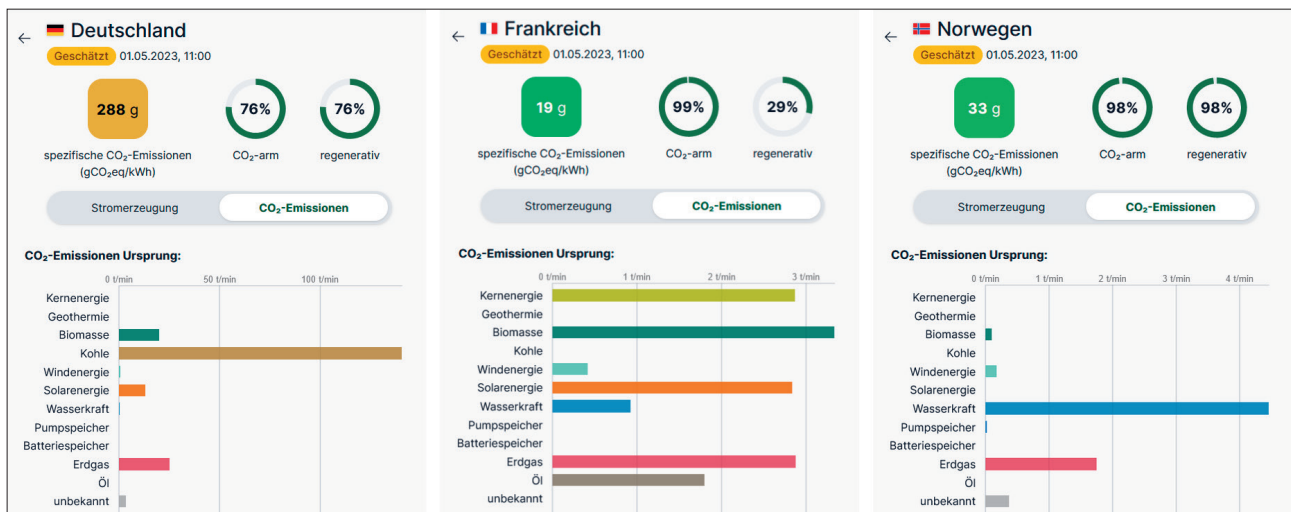
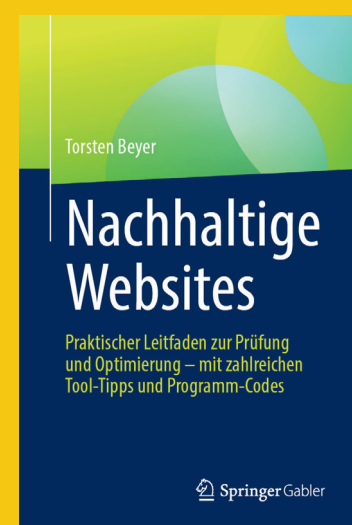


Abb. 6: Auf der Seite Electricity Maps (<https://app.electricitymaps.com/>) können fast in Echtzeit die Zusammensetzung des aktuell produzierten bzw. verbrauchten Stroms und die daraus resultierenden CO₂-Emissionen einzelner Länder eingesehen werden. Der schlechte Wert (288 g CO₂eq/kWh) für Deutschland resultiert aus der Verbrennung von zu viel Kohle und Erdgas. Frankreich (19 g CO₂eq/kWh) hat aktuell ein 15-mal grüneres Stromnetz, primär durch die Nutzung der problematischen Kernenergie. Bei regenerativen Energien liegen Länder wie Norwegen (33 g CO₂eq/kWh) vorn, dessen Netz 9-mal grüner ist. Man beachte die unterschiedliche Skalierung der Balkendiagramme!

die CO₂-Emissionen aus unterschiedlichen Gründen viel niedriger als in Deutschland sind. Nimmt man dann noch den Umrechnungsfaktor von 0,81 kWh Stromverbrauch pro Gigabyte Datentransfer dazu, der sich aus dem Quotient des gesamten Energieverbrauchs des Internets und dem gesamten Datentransfer auf Endnutzenseite ergibt, dann hat man den fehlenden Bezug zwischen Stromverbrauch, Datenvolumen und CO₂-Emissionen.

Mit dem globalen Mittelwert von 442 g CO₂e/kWh ergibt sich daraus ein Wert von 358 g CO₂-Emissionen pro Gigabyte übertragenem Datenvolumen. Damit kann man die ermittelten Werte aus den Logfiles multiplizieren, um eine Idee zu bekommen, in welcher Größenordnung die durch die eigene Website verursachten CO₂-Emissionen liegen. Die Modelle werden immer weiter verfeinert, aber es ist klar, dass Daten im Internet durch viele Länder mit sehr unterschiedlichem Strommix fließen und eine exakte Ermittlung des Stromverbrauchs für die Datenübertragungen aktuell ebenso noch nicht möglich ist wie die Messung des Stromverbrauchs einer Website in einem Rechenzentrum. Dazu kommt der große Einfluss des verwendeten Endgeräts

INFO



Wenn Sie tiefer in das Thema einsteigen wollen, sei das gerade erschienene Buch des Autors „Nachhaltige Websites – Praktischer Leitfaden zur Prüfung und Optimierung – mit zahlreichen Tool-Tipps und Programm-Codes“ als Einführung in dieses wichtige Zukunftsthema empfohlen. Torsten Beyer. ISBN: 978-3-658-410926. Springer Gabler 2023. 178 Seiten. 37,99 Euro (Taschenbuch), 29,99 Euro (Kindle).

und der Zugangsart (Glasfaser, Kupfer, 4G, 5G ...) auf den Stromverbrauch.

Dennoch bleiben solche Betrachtungen wichtig, um den Effekt der

Internetnutzung sicht- und greifbarer zu machen und ein besseres Verständnis für dieses unsichtbare Problem zu schaffen.

Fazit

Der Aufbau eines Monitorings zur Dokumentierung der Nachhaltigkeit einer Website auf Basis des Datenvolumens ist nur durch eine Auswertung der Server-Logfiles möglich. Denn dort wird der gesamte Datenverkehr bis auf eingebundene externe Elemente in Abhängigkeit von allen unterschiedlichen Konfigurationen erfasst. Die vorgestellten Metriken können eine nachhaltige Optimierung belegen und Fehlentwicklungen durch den Einbau nicht optimierter Dateien oder veralteter Formate aufdecken.

Einen CO₂-Wert pro Seitenaufruf oder pro Jahr zu nennen, der aus einer Einzelseitenanalyse abgeschätzt wird, klingt in der aktuellen Klimadiskussion vielleicht besser. Er kann mit den aktuell verfügbaren Tools allerdings nur sehr ungenau ermittelt werden. Daher sollte man sich auf das Datenvolumen fokussieren. Denn nur was man exakt messen kann, kann man monitoren und optimieren.